The Intelligent Content Framework

A practical approach to accelerating the Study Design and Regulatory Documentation Development Processes using a Rules-driven, Structured Content Authoring Solution Framework

Imperative: Improved efficiencies, consistency and quality of information through the clinical product development lifecycle.

Throughout the clinical development life cycle, thousands of documents are generated by life sciences organizations, many of which are submitted to regulatory agencies, partners or external public databases. Life sciences organizations have been applying structured authoring solutions for a number of years to support specific process needs that address narrow sets of clinical processes and standards, such as Product Information Management and Product Labeling. Additionally, regulatory requirements, such as the eCTD, have resulted in the implementation of solutions that support an XML backbone to dynamically assemble clinical content components from multiple sources into an XML structure for transformation and output to required submissions formats. As standards evolve or change, these solutions become expensive to modify and adopt to accommodate evolving standards.

Current Challenges:

While current solutions address the specific processes they have been implemented for, in order to achieve higher levels of efficiencies, consistency and quality of information across all clinical planning and documentation processes, a broader Topic based Structured Content Authoring framework solution approach is required to address the entire product development life cycle. Working with key business partners, Microsoft has defined a solution approach labeled the Intelligent Content Framework (ICF) to address this problem. ICF brings together a set of key disciplines and technologies to provide a holistic solution approach to address these key needs for the life sciences industry. Aligned with the principles of the Microsoft Connected Life Sciences Framework, ICF supports Topic Based information design disciplines and integrates best-of-breed technologies to support authoring, review/ approval workflows, and publishing for transformation and output of clinical documentation content for existing regulatory systems and public databases.

> In the absence of a broad-based, framework approach to address structured authoring across the clinical development life cycle, point solutions will evolve to address business needs. These point solutions will result in increased costs, fragmentation of information across multiple systems and repositories, increasing the complexity of maintaining consistency and quality of messages across documents.

Toward a Solution: ICF disciplines and tools improve efficiencies and quality for study design, planning and regulatory documentation production.

The Intelligent Content Framework for Regulated Industries is a transformational approach to Topic Based Structured Content Authoring, and provides a complete DITA-based (DITA is an open OASIS standard for structured content authoring and content re-use) XML Component Content Management Solution based on the following components:

- The platform is implemented using Microsoft SharePoint® 2010 and Office 2010® to support the enterprise content management, collaboration, authoring and workflow needs,
- DITA Exchange™: a complete DITA XML Component Content Management Solution by DITA Exchange A/S,
- InRule® irAuthor and InRule® Server, to support the rules development and rules engine for the automation of clinical process,
- SharePoint® compatible third-party add-on solutions, such as Nintex® or K2® workflow, to support the review/ approval and life cycle automation needs.

ICF will allow organizations to address key challenges they face in an incremental fashion across the study design, planning and regulatory documentation development life cycle, while reducing the costs and complexity of implementing point solutions to support key needs. ICF will allow organizations to realize the following benefits:

- 'Quality by Design': ensuring quality and consistency of information across all documents produced through the product development cycle

- Increased efficiencies through faster document production and approval cycles, and automated transfer to external systems
- Reuse of knowledge and experience across project teams
- Improved clinical decision support
- Transparency in research – protocol information to public tracking and regulatory databases, such as clinicaltrials.gov and clinicaltrialresults.org
- Significantly improving the overall 'time-to-document', which in turn will support the overall business goal for improving 'time-to-market'

Start with the ICF disciplines: Intelligent Content Design

Transforming the broader study design and clinical documentation processes requires approaching the solution by applying intelligent content design disciplines as well as implementation of the ICF technology platform for Structured Authoring. Organizations must address these non-technology disciplines as a prerequisite to the implementation of the ICF solution. The following paragraphs describe these disciplines and the ICF technology framework to support required functional and automation needs.

Business Processes

Current trial design and documentation business processes must be reviewed/ analyzed in order to define target processes to support the new Topic Based Structured Content Authoring paradigm. Due to the broad applicability of the solution, these processes can be addressed incrementally based on business priority for practical implementation, typically starting with the early phase studies The following study design and specific documentation production business processes need to be reviewed and designed for implementation.

- Study Design
- Document Creation/ Initiation
- Document Authoring
- Review/ Approval Processes
- Document Finishing and Assembly
- External Systems Integration – Information/ Document Export
- Automated content conversion to other required formats (e.g. for clinicaltrials.gov etc.) using DITA Exchange™ transformation and publishing services

1. Information Architecture and Design

Information Architecture and design for regulatory content entails a detailed analysis of the information within and across documents to:

- Define the content taxonomy/ metadata for Document and Topic components
- Define the Topic component-level document context (reused or re-purposed Topics)
- Usage analysis within and across clinical processes, products and documents,
- Identification of the component Topics defined at a practical level for authors,
- Authoritative data sources for data used within the component topics and documents.

ICF enables the creation of "interoperable" content components by utilizing SharePoint® Managed Metadata services to provide metadata and content taxonomy management. Specifically, the following analysis and definitions will be defined incrementally for documents as they are implemented within ICF:

- Information classification at the document and granular component topic levels. For example, a subset of classification:
  - Document:
    - Category: Content Management
      - Document Title
      - Created By
      - Creation Date
      - Last Modified Date
    - Category: Trial Entity
      - Study Number
      - Study Phase
    - Category: Product
      - Compound ID
      - Generic Name
      - Product Line

The information classification activities will leverage existing taxonomy and metadata specifications and extend them to support component-level specifications. This process also provides an opportunity to align to industry standards, such as CDISC ODM or BRIDG sub-domains and Topic extensions for the Structured Authoring program.

Separation of Content from Context

The topic-based approach allows for the separation of content from context. The topics are linked into applicable documents as reusable (non-modifiable) content, or

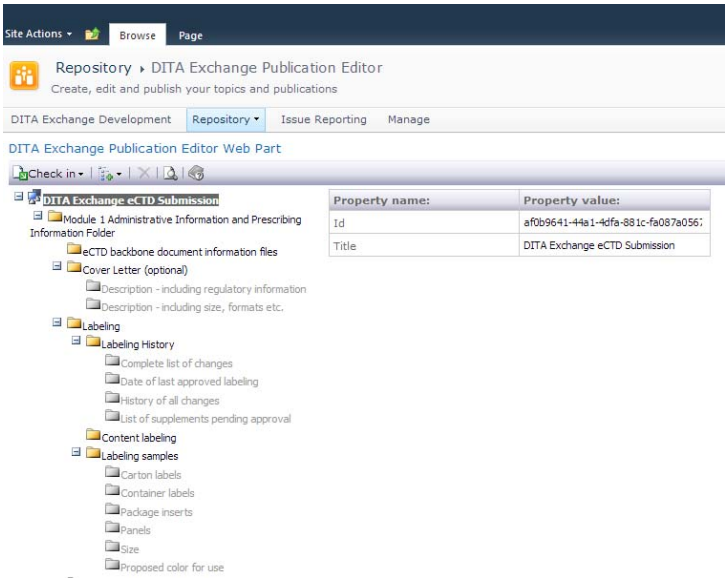
repurposed (modifiable) content, at which time the topic content has context within the document.

Document analysis and decomposition

Clinical documents are analyzed to decompose the document structure, and develop DITA document map templates that define the sections, topics and links to reusable or repurposed topics. The DITA Map can be viewed as the Table of Contents for the document, and provides the separation of content from context. During the author, review/ approval and document finalization processes, the map will be used to publish to multiple output formats, such as Word, PDF, HTML and various XML standards, including Open XML.

The document structures and a portfolio of documents for products will be defined and configured within the platform using the DITA Exchange™ Map Editor functionality. The portfolio map provides the means to create a master product documentation map, linking each individual document map as a sub-map within the master map. This approach allows reusable components linked into all clinical documents within the portfolio to be populated immediately at time of topic creation/ approval within the initial document, increasing production efficiencies.

The DITA Exchange™ Map Editor facilitates the creation, editing and management of the document and portfolio maps, as depicted below:



2. *Rules Design and Rules Authoring*

Throughout the study design and documentation development processes, a number of business rules drive decisions that determine document actions, such as the content that is included within clinical documentation, the security or other policies that are applied to the document at a specific life cycle stage. For example, criteria such as study phase, therapeutic area, and study type will drive the Inclusion/ Exclusion criteria included in the Synopsis, Protocol and other downstream clinical documents. While standard and optional topic content are created and managed within a Topics library, and linked into the document map template, when creating a new document or portfolio of documents for the product, applicable standard or optional component topic-content can be dynamically and automatically linked into the document instance created from the map template based upon specific business rules.

ICF supports the integration of a Microsoft® .NET based third-party rules-engine, InRule®, providing the tools necessary for the rules component as follows:

- Rules Design - a discipline followed as part of the information design process to analyze the state/ criteria for the business rules, and document corresponding actions
- Authoring of business rules using the InRule® irAuthor product
- InRule® Server to evaluate the rules and return the applicable state for application processing

Rules can be designed and authored for key processes such as:

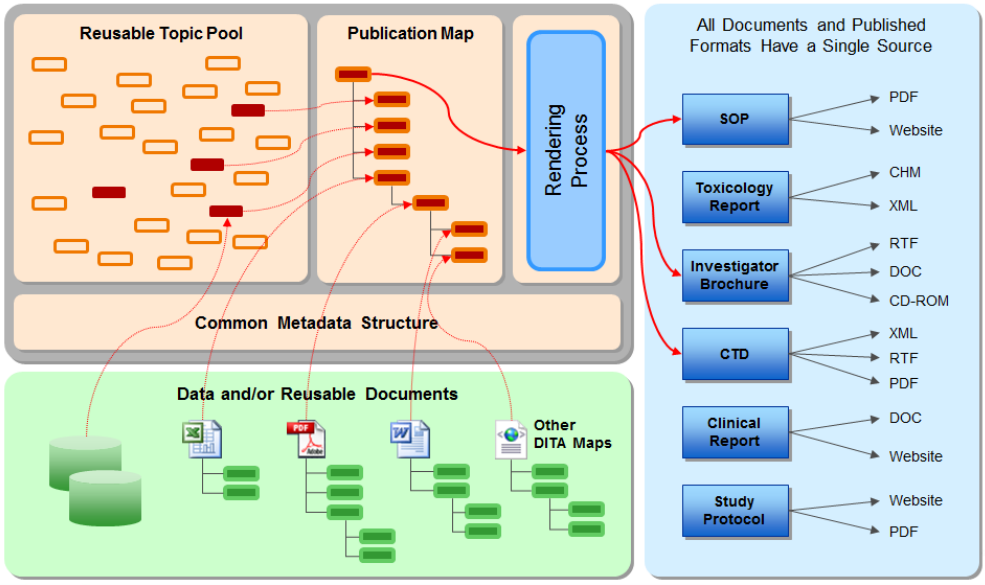
- Study Design Rules
- Document Initiation/ Creation Rules
- Review/ Approval Workflow Rules
- Document Finishing Rules
- Study Amendment Rules

The Rules engine is integrated within the framework, and will be invoked through Business Logic Services.

3. *Content Output Design*

Although DITA standards dictate that all content must be in XML, DITA Exchange™ provides a Word-based Open XML Publishing engine to allow organizations to “blend” traditional documents with reusable DITA XML topics, bringing a best-of-breed approach to support existing content and reducing transition risks.

The publishing engine can output the DITA XML file to multiple output formats using XSLT templates for transformation for required output, as illustrated below:

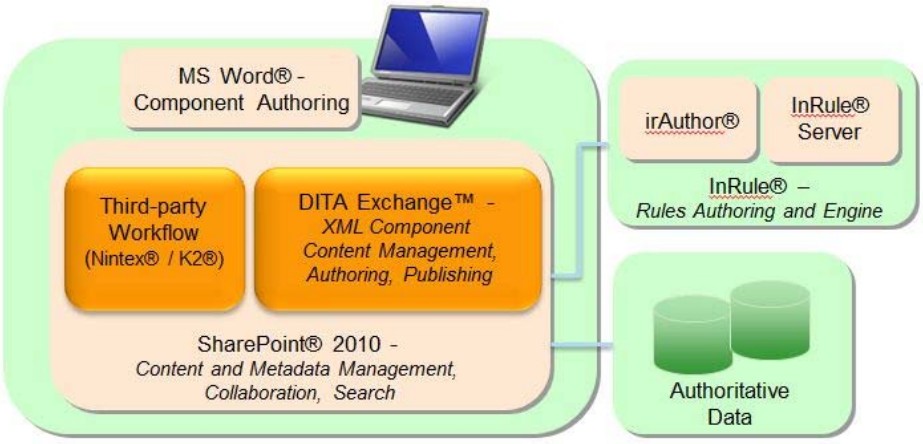


Authors must be trained to author the component topic-level content for related processes without necessarily being concerned about the context within the document. The new approach removes the burden of authors having to worry about format or non-value added activities.

Once the content is authored and approved, the Topics are automatically linked into the structure of one or more clinical documents that use this information based on the document map. As the content has already been approved through review/ approval cycles, organizations can realize improved efficiencies through the review/ approval cycle times since the content need not be re-approved.

4. Technology Framework

ICF integrates best-of-breed technologies, based on the Microsoft® .NET framework, Microsoft SharePoint® 2010 and Microsoft Office 2010®:



Change Management: New Roles and Disciplines

Moving to a Topic Based Structured Content Authoring paradigm requires users, such as medical writers and scientists who are used to authoring documents, to write content in a new paradigm. Authors will focus on authoring of the content, and approve the content once in the form of Topics. Current regulatory documentation processes must be designed to support this new paradigm.

The ICF approach introduces new roles and responsibilities, such as the role of the Information Architect. Information Architects will need to have sound business knowledge of the clinical documentation processes, and can be individuals who currently support the clinical and medical writers with the documentation services. They will require appropriate training in the design disciplines as well as the tools. The role of the Information Architect is to:

- Complete the content analysis and information design, working closely with the business
- Configure the managed metadata within the tool
- Configure the topic templates and document template maps
- Configure the Portfolio document map, as appropriate
- Configure the output style sheets for the publications process
- Support documentation services, such as document finalization and output to external systems and public databases

Organizations must plan change management aspects of the solution, such as process, skills and tools training.

ICF Benefits

The ICF solution provides a platform that can be deployed to incrementally realize the broad benefits of creating a "Content Supply Chain" strategy for clinical study documentation. Meeting diverse needs requires the incrementally implementation of a solution that is easily configurable and deployed based on business priorities that fits into the organization's culture.

ICF builds on four pillars:

- Built on standard Microsoft Office® and Microsoft SharePoint® Server platform for ease of use, system integration, scalability and cost efficiency
- The DITA XML standard supported by DITA Exchange™ as an Information Management Model to enable dynamic content reuse and to supplement the EDM Reference Model
- Integration of InRule® to support a rules-driven, process automation approach to improve efficiencies and quality of information
- Workflow integration, using a third-party workflow product such as Nintex® or K2®, to improve review/ approval processes, and automate component and document life cycle management

---

References

- Enterprise Content Management in Regulated Industries – Gabor Fari, Microsoft, June 2006
- Structured Authoring Best Practices – Vasu Ranganathan, ArborSys Group, 2009
- Intelligent Content Framework for Regulated Industries – Gabor Fari, Microsoft; James Averback, Life Sciences Integration Partners
- DITA Exchange™ and the Intelligent Content Framework [ICF] – Ole Rom Anderson, 2009
- Microsoft Connected Life Sciences Framework

For More Information

For more about the Microsoft Intelligent Content Framework, please contact Gabor Fari, Director, Life Sciences Solutions: Gabor.Fari@microsoft.com and please visit our Life Sciences Site at: www.microsoft.com/LifeSciences